基于发现系统的大学出版社开放学术资源现状调研及分析

武丽娜 左阳 窦天芳
清华大学图书馆 北京 100084

摘要:[目的/意义]调研了解大学出版社学术资源在发现系统中的开放现状和发展趋势。[方法/过程] 以两个主流发现系统 PRIMO 和 SUMMON 覆盖的 15 所大学出版社学术资源为对象，对这些学术资源在发现系统中的索引深度进行测试。统计发现系统中大学出版社学术资源的索引深度和索引规模，以及发现系统覆盖的大学出版社及其资源数量的变化趋势。以清华大学出版社为例，对国内大学出版社学术资源应加快开放共享进行探讨。[结果/结论] 发现系统中国外大学出版社的数量和学术资源数量呈现上升趋势，但订购刊和 OA 的索引深度未全部达到全文级别。中文资源的开放共享应受到关注。

关键词：大学出版社；学术资源；索引深度；
分类号： G250

作者简介：武丽娜（ORCID: 0000-0002-4955-7593），馆员，硕士，E-mail: wuln@lib.tsinghua.edu.cn；
左阳（ORCID: 0000-0002-6885-6417），馆员，硕士；窦天芳，副馆长，副研究馆员，硕士；

大学是知识生产中心，大学出版社为大学的教学科研服务，是大学与社会进行文化成果交流的桥梁[1]。随着“中国文化走出去”战略的部署和实施，面向国际社会传播中国知识文化成果，提升中国知识文化成果的国际影响力，成为大学及其所辖出版社、图书馆的共同责任。

开放是最好的知识传播方式。元数据交换及开放链接技术的成熟和应用是数字网络时代推进知识开放的核心力量。资源发现系统与出版社等内容提供商合作，通过元数据交换技术收集海量的、异构资源的元数据，通过预索引面向终端用户提供资源搜索服务，再通过开放链接技术实现最终的资源获取。截止到目前，国内外有超过7 000家出版社[2-3]与主流的资源发现系统合作，通过开放元数据甚至开放对象数据的方式提升资源的可见度。其中，大学出版社也在积极融入开放的学术传播生态系统，开拓各种渠道来提升传播影响力。Serials Solution 公司推出第一款网络级资源发现系统SUMMON以色列Ex Libris公司发布统一资源发现系统PRIMO作为两个主流的资源发现系统（这两个发现系统所在的两家公司已于2015年合并，但发现系统仍然独立运作）也收录了大量大学出版社的资源。例如，成立于1995年的斯坦福大学出版社运营的HighWire，其与学术出版社、学会、大学出版社等708家机构合作，收录3 000余种电子期刊、图书、会议录等资源，文章总数达7 659 003篇，其中免费文章达2 434 604篇[4]。2010年，英国牛津大学出版社将120万在线电子期刊文章移至HighWire 2.0平台；2013年，美国杜克大学出版社选择与HighWire合作建立新的电子书平台。HighWire已成为目前全球最大的提供免费全文的学术文献出版商，对以盈利为目的的传统商业性出版商形成了强有力的竞争[5]。成立于1534年的剑桥大学出版社出版360余种全文电子期刊，近72%的期刊被SCI/SSCI/A&HCI收录，30余种在学科内排名前10[6]。牛津大学出版社出版的电子期刊现有220多种同行评审期刊，其中60%的期刊是与学会及国际组织合作出版的，覆盖广泛的学术领域，其中有诸多全球被引率最高的学术期刊和文章[7]。而且，牛津大学出版社和剑桥大学出版社也先后于2004年和2006年开始实施开放获取项目[8]。可见，大学出版社正逐渐以更加开放的姿态，通过开展合作或开放获取的方式，推进科学成果的传播。

统一资源发现系统被当前大多数高校和研究机构采用，帮助用户一站式检索和获取资源。当前，图书馆界对发现系统的研究集中在不同发现平台的比较和元数据评估方面[9-10]，对发现系统中学术资源开放程度的研究未见报道。本文试图从大学出版社学术资源在发现系统中的索引深度和发现系统覆盖的大学出版社数量两方面进行调研，结合清华大学出版社开放获取期刊的实践，推测大学出版社学术资源在发现系统中的开放趋势。

1 知名大学出版社学术资源开放现状调研

1.1 知名大学出版社的遴选

原Proquest公司发现系统SUMMON覆盖了84家大学出版社[3]。清华大学图书馆自2011年引进Ex Libris公司发现系统PRIMO，其后台资源覆盖了剑桥大学出版社、杜克大学出版社、爱丁堡大学出版社、牛津大学出版社、芝加哥大学出版社和麻省理工大学出版社等20余家大学出版社。笔者根据世界大学排名[11]筛选出排名在前30位，以出版期刊为主，且被这两家发现系统收录的15家知名大学的出版社，通过网站测试的方式，分析这些出版社资源在PRIMO和SUMMON两个发现系统中元数据的开放现状。

这15家大学出版社包括Cambridge University Press（剑桥大学出版社）[12]、Coimbra University Press（科英布拉大学出版社）[13]、Cornell University Press（康奈尔大学出版社）[14]、Duke University Press（杜克大学出版社）[15]、Edinburgh University Press（爱丁堡大学出版社）[16]、Johns Hopkins University Press（约翰霍普金斯大学出版社）[17]、Massachusetts Institute of Technology（麻省理工学院出版社）[18]、Michigan State University Press（密歇根大学出版社）[19]、Oxford University Press（牛津大学出版社）[20]、Purdue University Press（普渡大学出版社）[21]、RMIT Publishing（皇家墨尔本理工大学出版社）[22]、University of California Press（加利福尼亚大学出版社）[23]、University of Chicago Press（芝加哥大学出版社）[24]、University of Toronto Press（多伦多大学出版社）[25]、University of Massachusetts Press（马萨诸塞大学出版社）[26]。这些大学出版社均为知名大学出版社，在学术界享有很高声誉。这些大学出版社的出版物学科范围和资源类型不尽相同，如剑桥大学出版社既出版期刊也出版图书，其期刊涉及人文社科、科学技术和医学学科；MUSE收录期刊的学科重点在文学、历史、地域研究等。再如，有些出版社如普林斯顿大学出版社、斯坦福大学出版社只出版图书。本文主要以这些出版社的期刊资源为调研重点。

1.2 发现系统中大学出版社学术资源索引深度调查及启示

发现系统元数据内容有相对而言的“薄、厚”之分。所谓“薄”元数据(thin metadata)是指包含字段非常少的元数据信息，包含题名、作者、合作者、来源、标识等基本信息；如一条记录不仅包含了上面提到的基本信息，还包含主题、摘要等信息，这就是典型的“厚”元数据(thick metadata)。“厚”元数据在基于题名、来源、作者等基本的检索中体现不出优势，但在数据分析、重组以及重用、数据挖掘等方面呈现出明显优势[27]。

本文主要以期刊文章数据为调研基础，以索引深度为测试对象，并以此得出其索引规模（全文、文摘、题录的比例）。具体方法为：在出版社官方网站随机选取三种不同期刊，分别在这三个期刊中随机选取一篇文章，先后提取篇名、文摘和正文的部分字段，分别在清华大学“水木搜索”平台（发现系统为PRIMO）和北京大学“未名搜索”平台（发现系统为SUMMON）进行检索测试，如根据篇名、文摘和正文部分字段均可以检索到该篇文章，则该文章所在期刊在发现系统中的索引深度为“全文”；如根据篇名、文摘部分字段可以检索到该篇文章，而根据正文字段不能检索到该篇文章，则该文章所在期刊在发现系统中的索引深度为“文摘”；如只能根据篇名检索到该篇文章，则索引深度为“题录”。对这15家出版社的测试结果如表1所示。表1中“索引深度”一项所列为测试中遇到的几种情况。

表1 大学出版社期刊资源在PRIMO和SUMMON中的索引深度调查结果

大学出版社	PRIMO		SUMMON	
	是否作为独立资源收录/我馆是否订购期刊资源	**索引深度**	**是否作为独立资源收录/我馆是否订购期刊资源**	**索引深度**
Cambridge University Press	收录/是	全文/文摘/题录	收录/是	全文/文摘/题录

Coimbra University Press	收录/否	全文/文摘/题录	未独立收录，收录在其他数据库中/否	文摘/题录
Cornell University Press	未独立收录，收录在其他数据库中/否	题录	收录/否	题录
Duke University Press	收录/否	全文/文摘	收录/否	全文/文摘
Edinburgh University Press	收录/否	全文/文摘/题录	收录/否	全文/文摘/题录
Johns Hopkins University Press	收录/是	全文/文摘/题录	收录/是	全文/文摘/题录
Massachusetts Institute of Technology	收录/否	全文	收录/否	全文
Michigan State University Press	未独立收录，收录在其他数据库中/否	文摘/题录	收录/否	文摘/题录
Oxford University Press	收录/是	文摘/题录	收录/是	文摘
Purdue University Press	收录/否	全文/文摘	收录/否	文摘
RMIT Publishing	收录/否	文摘	收录/否	全文
University of California Press	未独立收录，收录在其他数据库中/否	全文	收录/否	全文
University of Chicago Press	收录/否	全文/文摘	收录/否	全文/文摘
University of Toronto Press	未独立收录，收录在其他数据库中/否	文摘/题录	收录/否	文摘/题录
University of Massachusetts Press	未独立收录，收录在其他数据库中/否	全文/文摘	收录/否	全文/文摘

笔者根据调查结果得出如下几点结论：

（1）根据索引深度调查结果，统计出这些出版社期刊在发现系统中的索引规模（索引规模=（出现该索引深度的出版社数量 / 出版社总数量）*100%）这些出版社期刊资源在PRIMO中的全文索引规模达到了62.5%，文摘索引规模达到了75%；在SUMMON中的全文索引规模达到了56%，文摘索引规模达到了69%。详细结果如表2所示：

表2　大学出版社期刊资源在PRIMO和SUMMON中的索引深度调查结果汇总

发现系统	发现系统PRIMO	发现系统SUMMON
索引深度	索引规模（即达到该索引深度的出版社数量所占百分比）	索引规模（即达到该索引深度的出版社数量所占百分比）

全文	62.5%	56%
文摘	75%	69%
题录	50%	44%

（2）同一个出版社的期刊索引深度在发现系统平台的索引深度不一致。测试发现，同一出版社的同一种期刊在两个不同发现平台的索引深度不一致，或同一出版社的不同种期刊在同一个发现平台的索引深度不一致。索引深度的差异是因为这些文章同时被多个不同的数据库收录，而这些数据源向PRIMO或SUMMON提供的元数据薄厚不一，如数据源向PRIMO提供了全字段，PRIMO也做到了全字段索引，则索引深度可以到全字段，以此类推，这样就导致了索引深度的不同。

（3）订购刊的索引深度未全部达到全文级别。测试发现，订购数据库的期刊资源索引深度并未全部达到全文。例如，我馆订购了剑桥大学出版社、牛津大学出版社和约翰霍普斯金大学出版社的期刊资源，但是这三家出版社也并未开放所有期刊全文的元数据。如牛津大学出版社期刊的索引深度只能到文摘或题录级别。测试也发现一些大学出版社的期刊虽然被其他商业数据库收录，但是这些期刊的索引深度也受数据库开放政策的制约。

（4）开放获取的日趋发展促使学术期刊出版机构逐渐加入开放获取出版的阵营，并迅速发展。目前，包括大学出版社在内的国际很多知名学术期刊出版机构都实施了开放获取[28]。本次调研发现，牛津大学出版社、麻省理工学院出版社、普渡大学出版社等都有独立的OA刊平台，笔者也重点测试了这些大学出版社的OA期刊和free期刊，发现这些期刊的索引深度或为文摘，或为全文，并未达到全部为全文级别的理想状态。

（5）发现系统独立收录的大学出版社全部为国外大学出版社，没有收录国内知名大学出版社。原因一方面可能是由于平台开发商为国外公司，所以收录的大学出版社以国外知名大学为主，另一方面是由于国内知名大学出版社数字化的进展稍落后于国外。

（6）元数据的准确性有待提高。测试发现，个别截取的测试字段不能获得结果，但是另外一些测试字段可以获得结果。这可能是由于元数据不够准确，同样会对检索质量带来影响。

1.3 发现系统中大学出版社数量和学术资源数量变化趋势

笔者于2015年9月、2016年9月和2017年4月分别调研了PRIMO发现系统独立收录的大学出版社期刊资源情况，发现2016年独立收录的出版社数量为10家，2016年9月为12家，而2017年4月增长到21家（见图1）。这些出版社开放的资源也不仅限于期刊，还包括其他资源类型。如牛津大学出版社开放的资源包括了期刊、

Primo Central Index 覆盖的大学出版社数量

25
20
15
10
5
0

（10 家）
Cambridge Univ.
Cork Univ.
Duke Univ.
Edinburgh Univ.
MIT Univ.
Oxford Univ.
Johns Hopkins Univ.
California Univ.
Chicago Univ.
Wales Univ.

10

（12 家）Add2
Edinburgh Univ.
RMIT Publishing

12

19

（19 家）Add7
Amsterdam Univ.
Hamburg Univ.
Coimbra Univ.
Leuven Univ.
Northwestern Univ.
Purdue Univ.
State University of New York

2015 年 9 月
2016 年 9 月
2017 年 4 月

图书、报告、百科、参考资料等多种类型；而且其开放至PRIMO发现系统的元数据不断增加，每年都有新增的资源集合。牛津大学出版、普渡大学出版社将OA资源也开放给了发现系统。发现系统中大学出版社期刊资源的数量呈现上升趋势表明更多的学术出版机构愿意通过协议或者开放获取的方式开放元数据乃至全文信息，以期通过更广泛的检索将更多的用户带回到资源出版平台。

图1　2015-2017年PRIMO Central Index覆盖的大学出版社数量

2　清华大学出版社期刊开放现状

清华大学出版社期刊中心成立于2011年3月，目前出版22种期刊，被核心数据库收录的英文期刊数量一直保持上升空间，社科类期刊学术影响力也有了显著提高，已经形成既有英文又有中文、既有科技又有社科的学术期刊集群雏形。其中自主OA期刊14种， 合作OA期刊6种,这些开放获取期刊都有可靠的出版平台。

目前清华大学期刊中心的大部分期刊都建有自己的网站和投审稿平台，并与国内外大型文献数据库合作进行期刊推广。在此基础上，期刊中心建立了自己的门户网站（www.tsinghuajournals.com），集宣传和服务功能于一体，由专人负责管理维护。同时积极发挥新媒体优势，开通微信公众号(TUP_Journals)，及时发布期刊中心和各刊动态。网站和微信公众号相互配合，注重期刊中心的整体宣传效应，树立清华大学学术期刊的品牌[28]。

清华大学出版社依托清华大学的学科优势，充分发挥期刊出版的集群优势，不断提升期刊的学术影响力，推动中文资源国际化，并且利用国际国内两个市场、两种资源，增加版权输出数量，加大与各种版权组织的合作力度，积极与国际知名数字版权运营商合作，拓展数字版权输出的广度与深度。

在数字平台建设方面，清华社建立了数字出版门户网站——文泉书局等四个数字资源服务平台和新技术服务公司；同时加强官方网站、微信平台的规划与管理，积极探索传统出版向新媒体出版的转型之路。未来清华大学出版社将积极打造开放存取出版与发布平台，促进内容集聚和品牌建设，还将搭建“清华学术在线”，建设清华社学术图书与期刊统一数字化平台。同步推进内容建设和数字化建设步伐，通过数字化带动内容资源建设，实现学术期刊全流程、全方位网络出版[30]。

3　结论

通过对两个主流发现系统覆盖的15所大学出版社期刊资源在发现系统中的索引深度的抽样测试以及发现系统中大学出版社数量变化趋势的分析，可以得出：发现系统中国外大学出版社的数量和学术资源数量呈现上升趋势。但值得关注的是，订购刊和OA刊的索引深度未全部达到全文级别。大学出版社学术资源的开放共享，是大学出版社发展、用户需求驱动和发现系统赢得用户的必然趋势，有利于促进知识的传播和发现[31]。笔者认为，大学出版社作为资源的所有者，应在促进学术资源的开放和传播中发挥更加积极的主导作用。

在调研中也发现国内大学出版社的缺失，作为大学出版社，应该认识到数字出版对学术传播的巨大推动力，着眼于数字出版营利方式的探索和产品服务的开发，让出版社更好地适应新媒体、新技术下的学术传播和出版市场环境[32]。与此同时，大学出版社也应该从自身出发，努力提升旗下出版物的学术影响力，促进中文学术资源的国际化。要积极“走出去”参与国际合作，在增强吸引力、承载力、聚合力上下功夫，推动数字版权在国际国内市场的输出，以合作共赢推动更大发展。

通过对大学出版社在发现系统中学术资源开放情况的调研，笔者得到几点启示：

（1）坚持把提高期刊学术影响力作为出发点和落脚点，不断深化大学出版社的品牌建设，拓展大学出版社品牌外延，提高品牌价值，最终提升自身的核心竞争力。

（2）从目前的现状来看，数字化终将是学术出版未来的主流趋势。大学出版社可以抓住“互联网+”的契机，立体化、多方位地探索适合自身的转型融合之路。同时，在实践中应该同步推进内容建设和数字化建设步伐，通过数字化带动内容资源建设，将自身重点定位为教育内容和学术内容、专业内容的集成者、供应商和服务商，形成一个自主的数字产品开发与营运平台[33]。

（3）大学出版社要积极利用国际国内两个市场，积极探索与国际知名数字版权运营商的合作，促进中文学术资源的国际化，增强自身的吸引力、承载力以及聚合力，拓展数字版权输出的广度与深度。

随着数字化学术文献的广泛普及，学术文献的开放及共享渐成趋势，数据本身的规范以及支持数据交换的协议日趋成熟，相信会有更多的大学出版社愿意通过协议或者开放获取的方式开放元数据乃至全文信息，以期通过更广泛的检索将更多的用户带回到资源出版平台。同时，将学术资源开放获取理念与现代信息通信技术相结合，让流动的知识（数据）释放出来，成为共享资源，将促进学术成果的揭示、开放和再利用，形成有利于知识共享、大众创新和经济发展的科研环境，进而推动科技创新[34]。

参考文献：

[1]史永霞.浅析大学出版社品牌[J].改革与开放，2009(5):145-146.
[2]Primo Central Collection list [EB/OL].[2017-07-11].https://www.sheffield.ac.uk/library/cdfiles/pclist.
[3] Summon participating publishers[EB/OL].[2017-05-25].http://media2.proquest.com/documents/Report-Summon-Publishers.pdf.
[4]Highwire[EB/OL].[2017-07-13].http://home.highwire.org/.
[5]赵廓，彭远红，颜帅.HighWire 学术期刊出版平台功能研究[J].科技与出版，2013(11):103-106.
[6]Cambridge电子期刊[EB/OL].[2017-02-01].http://www.lib.tsinghua.edu.cn/database/cambridge.htm.
[7]Oxford电子期刊[EB/OL].[2017-02-01].http://www.lib.tsinghua.edu.cn/database/oxford.htm.
[8]刘晓霞，张新鹤.国际著名学术期刊出版机构的开放获取政策研究[J].出版发行研究，2016(9):57-62.
[9]程颖.资源发现系统元数据的问题与思考[J].图书情报工作，2015,59(9):104-110,126.
[10]秦鸿，钱国富，钟远薪.三种发现服务系统的比较研究[J].大学图书馆学报，2012,30(5):5-11.
[11]World University Rankings[EB/OL].[2015-09-01].https://www.timeshighereducation.com/world-university-rankings/2017/world-ranking#!/page/0/length/25/sort_by/scores_overall/sort_order/asc/cols/stats.
[12]Cambridge University Press[EB/OL].[2015-10-10].https://www.cambridge.org/core.
[13]Coimbra University Press[EB/OL].[2015-10-10].http://www.uc.pt/en/imprensa_uc/generalinformations.
[14]Cornell University Press[EB/OL].[2015-10-10].http://www.cornellpress.cornell.edu/.
[15]Duke University Press[EB/OL].[2015-10-10].https://www.dukeupress.edu/.
[16]Edinburgh University Press[EB/OL].[2015-10-10].http://www.euppublishing.com/journals.
[17]Johns Hopkins University Press[EB/OL].[2016-05-15].http://www.lib.tsinghua.edu.cn/database/muse.htm.
[18]Massachusetts Institute of Technology[EB/OL].[2016-05-15].https://mitpress.mit.edu/.
[19]Michigan State University Press[EB/OL].[2016-05-15].http://msupress.org/journals/.
[20]Oxford University Press[EB/OL].[2016-05-15].http://www.journals.uchicago.edu/.
[21]Purdue University Press[EB/OL].[2016-05-15].http://www.thepress.purdue.edu/.
[22]RMIT Publishing[EB/OL].[2016-05-15].http://www.informit.com.au/.
[23]University of California Press [EB/OL] . [2016-05-15].http://www.ucpress.edu/journals.php.
[24]University of Chicago Press[EB/OL].[2016-05-15].http://www.journals.uchicago.edu/.
[25]University of Toronto Press[EB/OL].[2016-05-15].http://www.utpjournals.com/.
[26]University of Massachusetts Press[EB/OL].[2016-05-15].http://www.umass.edu/umpress/.
[27]窦天芳，姜爱蓉.资源发现系统功能分析及应用前景[J].图书情报工作，2012,56(7):38-43.
[28]刘俊，张昕，颜帅.大学出版社学术期刊集群化运营模式研究——以清华大学出版社期刊中心为例[J].编辑学报，2016,28(6):561-565.

[30]柏荷.宗俊峰:践行清华精神担当文化企业的社会责任——再访清华大学出版社社长[J]. 出版广角，2017(6):30-34.
[31]雷鸣，裴琳琳.大学出版社数字出版的现实困境与路径选择[J].出版广角，2016(5):26-28.
[32]王跃.牛津大学出版社与剑桥大学出版社的数字出版研究[D].南京：南京大学，2014.
[33]宗俊峰.坚守使命 融合发展 重在实践——谈“互联网+”时代的出版人才培养[J]. 科技与出版，2016,(3):4-7.
[34]吴建中（上海图书馆馆长）：从开放获取到开放科学 [EB/OL].[2017-05-19].http://blog.sciencenet.cn/blog-213646-891225.html.

作者贡献说明：
武丽娜：内容调研分析、论文撰写和修改；
左阳：协助内容调研分析、论文撰写；
窦天芳：负责提出调研命题思路、论文审核修订。

Study on the Open Scholarly Resources of University Presses Based on the Discovery System

Wu Lina Zuo Yang Dou Tianfang
Tsinghua University Library, Beijing 100084

Abstract: **[Purpose/significance]** Exploring the quantity, the opening level and the tendency of scholarly resources of 15 university presses using PRIMO and SUMMON discovery systems. **[Method/process]** Examining the indexing depth and scale of journals of these university presses by the sampling test method. Tsinghua University Press was introduced as a case of opening and sharing scholarly resources. **[Result/conclusion]** The quantity of scholarly resources of university presses is increasing and their open tendency is evident. However, the indexing depth of the subscription and OA resources are not all at a full-text level. The opening of Chinese resources should be focused on.

Keywords: university press scholarly resources indexing depth open